DTIC FILE COPY

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) Same
6a NAME OF PERFORMING ORGANIZATION Dept of Computer & Info. Sci University of Oregon — 6b OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Office of Naval Research
6c. ADDRESS (City, State, and ZIP Code) Eugene, OR 97403-1202	7b ADDRESS (City, State, and ZIP Code) Code 1142PS 800 N. Quincy St. Arlington, VA 22217-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research — 8b. OFFICE SYMBOL (If applicable) Code1142PS	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N0001487K0321
8c. ADDRESS (City, State, and ZIP Code) 800 N. Quincy St. Arlington, VA 22217-5000	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61153N 42	RR042-09	RR042-09-02	4424206

11 TITLE (Include Security Classification)

Binocular depth and the perception of visual surfaces (unclassified)

12. PERSONAL AUTHOR(S)
Kent A. Stevens, Allen Brookes

13a. TYPE OF REPORT	13b TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM ____ TO ____	1988/9/27	28

16 SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Stereopsis, binocular vision, depth perception, vision

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

This technical report consists of two publications. The first is the first chapter of (Brookes, 1988), a Ph.D. dissertation completed this summer. This chapter provides an overview of the work on depth from stereopsis performed under ONR contract N00014-87-K-0321. The second part of the report is a reprint of an article appearing in Vision Research which describes a set of experiments that show an insensitivity to constant gradients of disparity and suggest that places with nonzero second derivitives of disparity are used for computing depth.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL John J. O'Hare	22b TELEPHONE (Include Area Code) (202) 696-4502 — 22c. OFFICE SYMBOL Code 1142PS

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

88 10 6 100

BINOCULAR DEPTH AND THE PERCEPTION OF VISUAL SURFACES

This technical report consists of two publications. The first is the first chapter of (Brookes, 1988), a Ph.D. dissertation completed this summer. This chapter provides an overview of the work on depth from stereopsis performed under ONR contract N00014-87-K-0321. The second part of the report is a reprint of an article appearing in Vision Research which describes a set of experiments that show an insensitivity to constant gradients of disparity and suggest that places with nonzero second derivitives of disparity are used for computing depth.

EXTRACTING PRIMITIVE SURFACE DESCRIPTIONS WITH STEREOPSIS
(Chapter 1 of a dissertation by Allen Brookes)

It has been known since the invention of the stereoscope by Wheatstone (1838) that differences in the positions of visual items from the different perspectives of the two eyes is a source of depth information. However, it is still not clear how these differences are translated into depth information or how this information is integrated with other sources of 3D information. Stereopsis is generally expected to produce the perception of distance from disparity. While stereopsis does appear to serve, in part, as a rangefinder, the perception of *surfaces* in depth does not correspond directly with the depth according to the disparities at each point of the surface. This dissertation presents evidence that when the disparities in an image suggest that a surface is present, the depth of points on that surface are not computed directly from disparities, but are reconstructed indirectly from properties of the detected surface such as curvature or discontinuities. Given this evidence it appears that when surfaces are present disparity is used primarily as a source of information for detecting surface properties such as discontinuities and curvature. That is, ***depth derives from a surface shape representation and not vice versa.*** In principle, there is much more information available in the stereo disparities than seems to be incorporated in the eventual 3D percept. However, this viewpoint suggests a parsimonious approach towards the integration of stereopsis with motion and other monocular sources of 3D information such as shading and contours, as I shall discuss.

The idea that depth derives from surface properties is a departure from accepted theories. Previous notions of the representations and processes involved in stereopsis are not adequate to explain this relationship between depth and surfaces. The long term goal of this research is discover what the correct representations and processes might be. The objective of this dissertation is not to completely describe these processes and representations but to describe a theory of stereopsis in terms of the strategies used by the visual system in deriving depth from stereopsis and in integrating stereopsis with other sources of 3D information. The main conjecture of this dissertation is that the primary strategy of stereopsis is to find regions that can be described as surfaces and then to use the descriptions of these surfaces for subsequent processing.

Below I describe this theory in as much detail as possible, given what is presently known. The bulk of this dissertation consists of empirical studies that led to the formulation of the theory and that offer support for many of the conjectures. I begin with a discussion of what I mean by depth and stereopsis and a discussion of how this work fits into the existing theories and empirical studies.

What is Depth and How Does it Relate to Stereopsis?

Apparent depth in an image is usually defined mathematically to be the difference of apparent observer distances between a given point and a reference point or distance (see, e.g. Foley 1980). Depth is related to distance, in that depth can be derived from differences of known distances. However, apparent depth is independent of apparent distance in that we can judge depth in situations in which we cannot judge distance. For example, when looking through lenses such as binoculars or a microscope, the surface variation is apparent but the distance is not.

Stereopsis, as a psychological term, simply means the perception of depth from stereoscopic images. The fundamental primitive of stereopsis is disparity, which is the angular discrepancy between the positions of a point in the two images. In principle, depth can be computed from disparity, where the relationship between depth and disparity is given by the following equation, assuming that D is much larger than d.

$$disparity = \left(\frac{I}{2}\right)\frac{d}{D^2}$$

Since the distance D is proportional to the angle of convergence of the eyes the depth d can be computed from disparity, I and the angle of convergence of the two eyes. Before we can compute depth, however, the individual points of each image need to be matched to produce the disparity values. This matching process has inspired a great deal of research, and the complexity of the problem has led to the belief that errors in the perception of depth from stereoscopic images result from errors in matching. Although incorrect matching can in fact have a considerable effect on perceived depth, there are also important cases in which the matching is unambiguous and yet the perceived depth is not predicted by the disparities. This decoupling of depth and disparity suggests that stereopsis is not simply a direct computation of a depth value for each point at which there is a disparity. Instead there must be some more global processes on which the perceived depth depends.

The principle concern in studying stereopsis is the horizontal disparity proportional to the depth to be derived. I refer to this as binocular or stereoscopic information even though the description includes some information that is binocular in the sense that there is presented to both eyes but does not present horizontal disparities. Equivalently, I use the word monocular to refer to 3D information that can be derived independent of whether it is presented to one eye or both. Thus, an image with right and left half-images contains a binocular component, the disparities, and a monocular component, which may include contours, occlusion, shading or a variety of other information that suggests that suggests surface relief.

Background

The focus of this dissertation is on how depth is derived from binocular disparities. In particular I focus on deriving depth from disparities for points associated with continuous surfaces. This area has not been explored theoretically before since there has been the more or less tacit assumption that depth was computed directly from binocular disparity. Thus the bulk of work in the area of binocular processing has been in the area of determining the correspondence between the left and right eye images. This dissertation starts with the assumption that the left and right images have been correctly matched and asks how depth derives from the resulting disparities. Deriving depth from disparity has only been studied from the point of view of finding geometrical constraints that would allow a direct computation of depth from disparities (Foley, 1980; Mahyew & Longuet-Higgins, 1982; Ritter, 1979). For conditions in which points are isolated there are results which accurately predict the perceived depth associated with particular disparities. However, the perceived depth of points on surfaces is not, in general, directly related to the disparities of those points. Since, until very recently, this has been virtually ignored in the literature I find much of the previous work to be irrelevant. Of relevance to this dissertation are isolated phenomenological studies that have shown instances in which the depth percept does not seem to derive directly from the disparities.

Recently, studies have emerged that throw some doubt on the previously accepted direct depth theory. Gillam *et al.* (1984) found that the presence of discontinuities in disparity reduces the time course of the development of, and increases the vividness of, the depth percept. This indicates that there is no simple conversion from disparities to depth since the conversion is presumably based on the disparity values themselves rather than the differences of disparities. Mitchison and Westheimer (1984) showed that for judgments of the relative depth of two lines, the presence of additional lines will change the percept. Additional lines lying in the same plane seem to increase the threshold for determining whether the two lines are at different depths. The result is that the lines are seen as lying on a plane parallel to the frontal plane, that is, the plane parallel to the plane containing the eyes. Generally, the depth interpretation of disparity requires that the stimulus present local disparity differences or contrast (Gogel, 1956, 1972; Gulick & Lawson, 1976). Contrast can incorrectly attribute relative depth to particular features as demonstrated in the so-called "depth contrast" effect (Werner 1938, 1942; Pastore, 1964; Pastore & Terwilliger, 1966). In the case of depth contrast effects, slant in depth can be induced in objects that have no disparity variation. This is done by contrasting these objects with objects with significant disparity variation. Ogle (1946) suggested that cyclotorsion (rotation of the eyes) in bringing the context to zero disparity, could change the disparities to fit the depth percept. Nelson (1977) later described various experiments that ruled out cyclotorsion as the sole explanation. Werner's (1938) primary observation, furthered by Nelson (1977), is that disparity *contrast* is responsible for the induction of apparent depth. That is, differences of disparity are more reliably related to depth in certain stereograms than are the absolute disparities. Mitchison and McKee (1985) showed stimuli in which the depth percept was an interpolation of depths at the edges of the figure. They found that in these stimuli the interpolation only occurred when the spacing of the dots was less than about 6′. That is, depth seems to be computed differently when the points are separated then when they are close together.

The results described above show that depth cannot be a simple function of disparity but must incorporate more global processes. Julesz (1978) introduced the notion of global stereopsis to resolve ambiguities in the local matching processes. The term global stereopsis, or globality, is used to mean that finding matches for points or features from the two eye images when there is some ambiguity depends not only on the possible choices of point or feature at the location where the match is taking place, but also on other points that presumably have some connection to that point. For example, points which have the same disparity as a possible match for a point with an ambiguous disparity may influence the matching process to choose that disparity. However, this notion of globality only concerns matching, so that once the ambiguities are resolved depth is assumed to be derived directly from the resulting disparities. Nelson (1975) likewise restricted this notion of globality by tying it to processes of facilitation and inhibition of disparity detectors. In this scheme the matching process is restricted to single matches by inhibiting other matches in the same visual direction and facilitating or strengthening matches with identical disparities. His model does not incorporate the integration of other 3D sources in computing depth however.

Stereopsis is one of many sources of 3D information that contributes to the eventual perception of space. There must be some integration of these sources into a single representation to form this percept. A clear example that integration takes place is the fact that although a monocular image may seem to have vivid relief, the addition of binocular disparities consistent with the monocular information gives a much more vivid impression of the relief. Equivalently, the combination is more vivid than that given by stereo alone. Even though this dissertation is mainly concerned with depth from purely binocular information, I find results concerned with the integration of 3D sources are also relevant since the goals of the computation of stereosis are affected by the need to integrate these other sources.

Studying integration may also offer clues about the representation. Richards (1977) provides evidence for one type of integration of multiple 3D sources. He reported a dramatic difference in the perception of depth in short (200 msec) presentations between random dot stereograms and equivalent stereograms containing monocular edges, and suggested that disparity discontinuities are most reliably interpreted when associated with monocular features. He further suggested that monocular cues act as a seed to the process of stereopsis. A related result is that of Gillam (1968) in which perspective was brought into conflict with disparity. The result in most cases was a compromise between perspective and disparity. In both cases the depth percept is affected by monocular surface information. For conflicts between depth from motion and disparity, Braunstein et. al. (1986) showed that in many cases the monocular interpretation dominates. Dosher *et al.* (1986) compared stereopsis and proximity luminance covariance in agreement and in conflict to find their relative strengths. The task consisted of determining the orientation or direction of rotation of a wire frame cube. They found evidence that the strengths of the individual cues were algebraicly added. Another type of evidence for integration is that motion parallax causes aftereffects in stereoscopic images (Rogers & Graham, 1984). Prolonged stimulation with a moving field of dots for which the motion path was consistent with a corrugated surface, tended to cause a flat stereo surface to be seen as corrugated with the opposite phase of motion image. If there were no integration between motion and stereopsis one would not expect such effects. Epstein (1973) addressed the issue of combining multiple sources of information into a single percept which he calls "taking-into-account". He points out that in some cases single sources of information do not have enough information to specify the percept and there must be an integration to provide the missing information. Epstein also provided a process model for the integration of multiple cues into a single percept and discusses various different perceptual examples where he believes that such an integration takes place.

Direct Depth vs. Reconstructive Depth

There is ample evidence that the human stereo system can accurately compute distances within about 2m (Foley & Richards, 1972; Wallach & Zuckerman, 1963; Ritter, 1977, 1979; Morrison & Whiteside, 1984), and, within that range perceived distance intervals, or depth, is directly related to disparity. It is reasonable to conclude that the human stereo system is predominantly a range finder which provides distances to each point in the image. It is also reasonable to suppose that the three-dimensional properties of surfaces are likewise derived from this range, and from local depth information. This conjecture is at least tacitly assumed in most computational models of stereopsis. These models can be summarized as follows:

$$\text{stereo disparity} \rightarrow \text{depth} \rightarrow \text{surface shape descriptors.}$$

In (Stevens & Brookes, 1987a) we call this model the ***direct depth model*** since depth is computed directly from the disparity values. Somewhere in this model must be incorporated a way to integrate monocular 3D cues. There are several alternatives. Stereopsis can be regarded as the dominant source of 3D information, specifically of depth, with other cues converted into depth. This information can be used to augment stereopsis where the sources are in agreement and to supplement stereopsis in places where stereopsis gives no information, e.g. out of range. This would imply that whenever stereo information was available and within range that the apparent depth would correspond with stereo disparity. This is not always the case, however, as will be examined in detail below. An alternative to this scheme is that other 3D sources are not always subservient to stereopsis and may override disparity information when there is conflict. Later it will be shown that there are cases in which there are no 3D cues other than stereopsis and yet the depth does not correspond to disparity. Many of these effects have in common that they are related to properties of surfaces. If these surface properties are computed first and then depth computed subsequently, one would expect artifacts related to reconstructing depth from surfaces. An alternative model that accounts for this lack of correspondence can be summarized with the following:

$$\text{stereo disparity} \rightarrow \text{surface shape descriptors} \rightarrow \text{depth.}$$

In (Stevens & Brookes, 1987a) we call this second model, the ***reconstructive depth model,*** since depth is determined from local "disparity contrast". The surface shape descriptors seem to consist of loci where disparity indicates a surface curvature or discontinuity feature. The determination of depth from shape features in stereopsis may be an evolutionary adaptation that allows stereo information about surface shape to be integrated with information from other sources. Thus other types of 3D information can be incorporated into the model in the following way:

$$\left.\begin{array}{l}\text{stereo disparity}\\ \text{shading}\\ \text{contours}\\ \text{motion}\\ \textit{etc.}\end{array}\right\} \rightarrow \text{surface shape descriptors} \rightarrow \text{depth.}$$

It seems much more feasible to combine and reconcile 3D evidence in terms of assertions about surface shape rather than primitive depth since, for most monocular 3D cues, properties such as surface curvature and orientation are more directly recoverable than object relief. It would be parsimonious, therefore to defer the computation of a depth map until the surface shape is decided.

The reconstructive depth model proposes that the goal of stereoscopic processing is to describe the visible surfaces in terms of their detected features and then to integrate this information with similar information provided from monocular sources. The following section describes a computational theory for how depth is derived from binocular information. The theory addresses monocular 3D information only to the extent that it is asserted that surface descriptions are the medium for integration of these sources with binocular information.

Computational Theory for the Reconstructive Depth Model

In Marr's view a computational theory is a description of the goals of the computation and the logic of the strategies for carrying out that goal (Marr, 1982). This differs from the notion of an algorithm in that the computational theory does not specify in detail how the steps of the computation are to be done. Marr (1982) states that the importance of a computational model is that it offers a high-level way of understanding a complex computation without having to understand low level implementation strategies that may be irrelevant to the goals of the computation.

In this section I develop a computational theory for stereopsis. The main goal of stereopsis, by this theory, is to find regions that can be considered to be surfaces and then describe to them. Much of how this is done is still not known, however this theory provides a framework that answers some questions about the main goals and strategies of stereopsis and provides questions for future work in this area.

Detecting Surfaces and Surface Properties

The first step in the reconstructive depth model is to find those areas that can be considered surfaces. The qualities of a surface defined only by binocular information depend very heavily on such parameters as the density of the points defining the surface and the presence or absence of points not associated with the surface. A particular set of points may or may not be seen as a surface if there are too many points in the same region that are not part of the surface. Also, if the number of points defining the surface is too small, they will appear as isolated points and not as a continuous surface. One way to discover why this is so is to look at how the visual system computes binocular disparities in the first place.

A certain percentage of the cells in the cortex are binocularly driven. The receptive fields of these cells for each eye are of similar size and orientation and are arranged in positions corresponding to a particular disparity (Hubel & Wiesel, 1962; Barlow, Blakemore, & Pettigrew, 1967). These cells will fire if both of their receptive fields are stimulated sufficiently and thus these cells have been called disparity detectors. The firing of a binocular cell is not really equivalent to detecting disparity, however, since there are instances in which the cell will fire without a point at that disparity. One instance of this is when there is a pair of points with the same vertical location but with different horizontal locations. These points will stimulate not only the receptive fields for the correct disparities but also the receptive fields for the disparity equal to the separation of the points. These anomalous disparities are not seen and can therefore not be considered detected. Since they are not seen despite the fact that their receptive fields are stimulated, the disparities must have been suppressed or ignored, allowing the correct percept to emerge.

There is, as yet, no neurophysiological evidence for how this suppression is done. One can, however, give a plausible explanation of how this is accomplished. Since the conflicting points would appear in the same visual direction the suppression could be done by having the strongest disparity signal for each visual direction inhibit any other disparity signals for that direction. The strength of a disparity signal could depend on various things. The cells corresponding to zero disparity seem to have greater numbers and strength so that with competition between zero and nonzero disparities the zero disparity percept should be seen. The fact that, for a surface, the places between points are seen as being on the surface suggests that adjacent points affect the disparity signals. This may be done by contributing strength to, or facilitating, adjacent disparity sensitive cell for the same or similar disparities. These facilitation and inhibition processes are similar to those suggested by Nelson (1975) for establishing binocular correspondence. A surface then may simply be an area which has a strong signal for a particular disparity. This corresponds well to the result that for a sufficient density of points the interstices are included in the surface.

Julesz (1971) states that for a sufficient density of points that the impression was no longer of a set of coplanar isolated points but of a solid surface covered with dots.

There are several possibilities of what to do with points that have disparities different than those of the surface points. One possibility is that when the surface has sufficient strength the points will be seen as lying on the surface. Another is that the perceived surface will not be seen as solid at that point and a discontinuity will be perceived associated with the point. What actually happens depends on the number of such points, how much the disparities differ from that of the surface, and the density of the surface. Again, a surface seems to be an area in which there is sufficient strength for a particular disparity, but in this case it is the strength as compared to the strengths of signals at other disparities. The competing disparity signals have less effect as the disparity difference increases.

Surfaces are Single Descriptions for an Area of Locations

Once a surface is detected a description must be formed. This need not be a strictly sequential process. The actual neural implementation may form the description at the same time the surface is being detected. Here the concern is only that a description of the surface is necessary. Each surface has a single description that precludes individual descriptions of the component elements of the surface. The converse of the hypothesis that an area seen as a surface has a single description is that when such an area is not seen as a surface, there is no single description for that area; each point is represented separately. Since each isolated point has a separate representation, accurate comparisons of their distances can be made. On the other hand, an area of points that is collected into a surface has a single representation and thus no longer has the properties of the individual points. In particular, I propose that there is no longer a depth value associated with each point. Instead a single overall distance value is associated with the surface. To get depth from the surface, or individual distances for points on the surface, it must be computed from the distance to the surface and other surface properties. Other properties are detected from the collection of surface points that are assembled into a description of the surface. These two properties are the discontinuities between surfaces (which are the edges of each surface) and the extremum points within the surfaces. From these edges and extrema the orientation of the surface is determined and the depths across the surface are reconstructed. The reconstruction, then, would consist of computing the depth of surface features and, when required, inferring the depth of a point using the assumption that there is a continuous surface between these features.

Detecting Properties of Surfaces and Reconstructing Depth

The reconstruction of depth from surface features is, in many ways, analogous to the reconstruction of brightness from detected changes in luminance features. Luminance changes, rather than absolute luminances, are detected. This provides for adaptability to a large range of luminances. Areas without detectable luminance changes are identified by correlating the borders where the changes occur. The brightness or perceived lightness of these areas is then reconstructed from the magnitudes of the changes along the borders of the region. The analogy holds to the extent that where surfaces are present, only changes in disparities are detected, and the depth of points within these surfaces are reconstructed from these detected changes. The detection of changes in luminance is accomplished by retinal ganglion cells which have a central excitatory region that sums the luminance within that area, and a surrounding inhibitory area which reduces the signal of the excitatory region by the sum of the surrounding luminance. These cells respond maximally when the center is filled with light and the surround is filled with dark. In order to then regain the correct luminance values, the inverse to this operation must be performed to reconstruct the lightnesses in places where there is no contrast. However, this reconstruction is not perfect and information is lost. As a result, a number of illusory brightness effects can be related directly to center-surround receptive fields. I will show later that there are no analogous effects for depth from disparity. This lack of analogous effects for depth suggests the possibility that there are no analogous center surround operators for depth. Another possibility is that the effects of these operators are nullified by the surface reconstruction process. For isolated points there do seem to be instances of effects similar to those due to center-surround receptive fields for luminance. Thus, I conjecture that disparity variations are measured by some sort of lateral inhibitory mechanism and are thus sensitive to disparity contrast and disparity curvature. For continuous surfaces disparity variation is slow so that regions without explicitly detected features will be presumed to be flat. This will effectively eliminate features induced by lateral inhibitory operators for regions with strong surface assertions.

How the reconstruction is accomplished is not clear. One method would be to interpolate depth values between features. This is not all that is done since essentially featureless planes can be seen as having a slant in depth, although underestimated. One possibility is that the interpolation can be augmented with attentive ranging information that would show some variation in a plane.

Integrating 3D Information

An important conjecture of the theory is that the property of "surfaceness" is a primitive in the representation of 3D objects from stereopsis. That is, the representation of objects consists, in part, of descriptions of the constituent surfaces. These descriptions are in terms of boundaries, surface orientation, curvature and possibly other features. Stereopsis is not the only source of information that contributes to the perception of 3D. Other sources of 3D information include shading, motion, and monocular contours. The final 3D percept is an integration of the information available from each of these sources. The evidence from the experiments presented here indicates that this integration takes place mainly at those places where there is surface information, that is, at the surface features of discontinuity and curvature. When there is agreement between sources the agreement strengthens the percept and thus creates a more vivid impression of depth. When there is disagreement between two or more sources the percept depends on the relative strength of the conflicting percepts and the constraints imposed by other features in the image. When the conflict is minimal (i.e., different degrees of curvature in the same direction) the percept is a compromise between the conflicting sources. When the sources suggest very different images then one source may dominate completely.

In natural images there is rarely disagreement between sources of 3D information. More often one or more sources will be ambiguous or have no information to offer. For example, there are many instances when part of the view to one eye will be obscured so that in that region stereopsis cannot occur. Yet we still get the impression of depth in these areas. In these cases the other sources fill in the information in a way that is consistent with the constraints imposed by the surrounding areas.

Summary of Supporting Results

The remainder of the dissertation is an attempt to verify that the claims made in the computational theory are correct. The proof rests on a set of empirical results that attempt to answer particular questions about deriving depth from binocular images. The next chapter discusses the methodology used in examining these questions and discusses why this is a reasonable approach. The following chapters discuss five sets of results individually. The following is a brief description of each set of results, the experiments involved, and the relevance of the results to the dissertation.

Depth from Monocular Contours is Commensurate with Depth from Stereopsis

A prerequisite of combining sources of 3D information is that they be in the same form. This is not to say that they all produce depth or that they all produce surface orientation. It is also possible that computations can be readily performed on one representation to produce the other. This issue of how to represent 3D information must be explored with a different paradigm.

In chapter III I describe several experiments in which the task was to judge the depth of a stereo probe point in relation to a monocularly presented surface. The surface in each case was a sinusoidal surface rendered with contours slanting back in space. The surface was presented in both perspective and orthographic projections. The judgments were made at four equally-spaced probe locations along a straight line on the surface, parallel to the ridges and troughs of the sinusoidal surface. The experiments differed in the presentation times and in the presence or absence of a stereo fixation. The results are also available in Stevens and Brookes (1987b).

In each experiment the resulting depth measurements showed monocularly increasing depth along the four probe locations. The two central probe locations provided a steep gradient in depth that showed no significant difference across the experiments. Three general conclusions were drawn from these experiments. First, depth is derived from both orthographic and perspective projections as a scaled quantity that is commensurate with the depth perceived from stereopsis in the near field. Second, the comparison of monocular and stereo depth is rather fast (achievable in exposures of only 150 msec) and does not require eye movements. The third conclusion is that the absolute distance to a fixated monocular surface is assumed to coincide with the stereo horopter, the set of points that have zero disparity. Binocular vision generally puts a fixated surface point in sharp focus and at zero disparity. Likewise, a fixated surface point in a monocular image, seen in sharp focus, is apparently regarded as lying at the same absolute distance as it would be if viewed binocularly at zero disparity.

Depth from Conflicting Monocular and Stereo Sources

The experiments above establish that 3D information from monocular and stereo sources can be compared. In general I would like to know how the visual system integrates the 3D information derived from stereopsis with that derived from other sources. The experiments described below suggest that the visual system does not reconcile certain types of conflict between the 3D information implicit in the stereo disparities, and the 3D interpretation derived monocularly. The findings might suggest the rivalry between monocular and stereo interpretations are often resolved in favor of the monocular, but since this is not the case for all stimuli and subjects, a preferable interpretation is that certain types of disparity gradient information are not processed, and the monocular interpretation was taken in the absence of detected information to the contrary. In either case, the results argue against certain earlier proposals for depth integration that otherwise seem intuitive, attractive, and computationally well-founded.

A series of experiments are presented in chapter IV (see also Stevens & Brookes, 1988) to attempt to determine what role stereopsis plays in the presence of contradictory monocular information. Experiment 1 concerned whether stereopsis could be used to effectively contradict the monocular interpretation of oblique intersections as foreshortened right angles, when the intersections were actually not perpendicular in 3D. The stimuli were planar grids and pairs of crossed lines in which the lines intersected at 90, 105, 120 or 135 degrees. Monocularly, this skew could be interpreted as a different slant to the plane in which the grid or cross is embedded. The task was to judge whether the intersection was skewed. It was found that stereopsis is remarkably impotent in influencing the perceived orientation and 3D configuration especially with the grids. Experiment 2 similarly examined relative depth judgments in displays with conflicting stereo and monocular information. Given a simple pair of stereo points, that with the greater (more positive) disparity is seen as relatively farther. But if these points are embedded in a continuous 3D surface, and if the monocular interpretation suggests an alternative relative depth between the two points, that monocular interpretation governed the judgment in the experiment. Experiment 3 similarly examined whether a conflicting disparity gradient influenced the monocularly interpreted surface orientation.

In these experiments the stimuli consisted of planar surfaces in 3D. Examination of control stimuli indicated that sufficient stereo information was available. Thus stereo disparities across a planar surface are not effectively analyzed in 3D. More formally, we hypothesized that stereopsis extracts 3D surface information only where the second spatial derivatives of dispartity are nonzero, corresponding to loci where the surface is curved, creased, or discontinuous. Experiment 4 directly examined planar versus nonplanar stereo stimuli, with and without competing monocular interpretations. The results further support this hypothesis. (And reviewing earlier studies, we observed that where stereopsis was particularly ineffective against conflicting monocular information, those studies also involved planar surfaces.)

These results suggest that depth is derived from disparity only where the surface exhibits continuous curvature or sharp discontinuities. Also, depth is reconstructed from multiple sources of evidence about surface topography. That is, surface shape is first analyzed in terms of sharp edges and creases, smooth folds, indentations, and so forth, from both binocular and monocular sources. The depth one experiences is a consequence of how this information is interpreted and reconciled. Depending on how the monocular information is interpreted, radically different depth distributions might be experienced. This is quite distinct from the notion that depth (and slant) is derived directly from stereo disparity (and its gradient).

The Effects of Surfaces in RDS

The results described in chapter V establish the major conjecture of this dissertation. That is, binocular depth is computed subsequent to computing surfaces and that depth is computed from the surface descriptions. In establish this conjecture I show that the depth of points can be influenced by the presence or absence of a surface. Since I am concerned with purely binocular depth, a continuous surface consists of points of a random dot stereogram in which the disparities are consistent with those of a particular continuous surface.

An experiment was performed to test this conjecture. The stimulus was a random dot stereogram with two different configurations. The first consisted of four slanted panels arranged roughly in a stairstep pattern. The slants of the panels were such that each panel had points of greater or lesser disparity than points on each other panel and yet had the overall impression of a set of slanted stairsteps. The other stimulus consisted of the same locations as the dots of the first stimulus but the disparities were randomized so that the disparity of each point was somewhere within the range of disparities of the first stimulus. The task, in the case of the paneled stimulus, consisted of showing one of the stimuli with a pair of probe points either on adjacent panels or on the outer pair of panels. For the random stimulus the same disparities were used which placed the probe points within the volume in depth. The subject was to decide which of the probe points was closer to the subject. The probe positions consisted of points that had equal disparities, points with greater disparities than those further up the stairsteps, and points with lesser disparities than further up the stairsteps.

The results of this experiment showed a significant effect in depth judgments between the surfaces and the random stimulus. For the random stimulus the pairs of probe points with different disparities were seen almost entirely correctly. Those with equal disparities elicited about equal judgments of nearer and farther indicating that they were also seen correctly. For the surface stimulus, the judgments for the probe points on the separated panels were consistent with a stairstep with little or no slant. This indicates an underestimation of the slants of the panels. For the adjacent panels, the depth of the probe points with larger disparity differences was judged correctly, but judgments for the probe points with smaller disparity differences and those with equal disparities again seemingly indicated underestimations in the slant of the panels.

If the depth of the pair of probe points were determined by a direct comparison of the disparities then the disparities of adjacent points should not effect the judgment. It appears that adjacent points which do not provide evidence of a surface do not effect the judgment. When the adjacent points are consistent with a surface, however, the judgment seems to be consistent with the properties of the perceived surface. This not only shows that the depth is derived from the surface but also adds support to the conjecture that surface properties such as slant are inaccurately derived from disparities. These results are discussed more thoroughly in Brookes and

Stevens (1988b).

Depth is Analogous to Brightness

Another major conjecture of the dissertation is that depth is a reconstructed quantity for non-isolated binocular points. This reconstruction seems to be based on places in the image in which the second derivative is non-zero. These places, which include discontinuities and curvature features, were earlier found to be important in processing disparity information. Analogously, in the luminance domain, it has been established that there are mechanisms sensitive to discontinuities and extrema of luminance. Various contrast illusions in the luminance domain have counterparts in the disparity domain with similar behaviors. These facts suggested that depth might be processed in a manner similar to brightness.

Chapter VI explores this analogy by comparing known brightness illusions with their depth counterparts. Much work has been done with brightness, and the underlying mechanisms responsible for this processing are fairly well understood. Since only changes in luminance are detected, perceived brightness is largely a reconstructed quantity. The mechanisms involved in the detection of luminance differences induce lateral inhibition effects which take the form of illusory bands or spots at areas of changing contrast. If brightness and depth were completely analogous, depth would show some type of lateral inhibition effects as well as reconstruction effects.

Various types of illusions were compared to test specific parts of the analogy. Patterns were used that are directly analogous to patterns which exhibit brightness contrast effects in the luminance domain. Changes in luminance were mapped to changes in disparity. It was discovered that illusions due to reconstruction of brightness values have counterparts in depth perception but that those due to spatial lateral inhibition do not. These results are also presented in Brookes and Stevens (1988a).

Detecting Surfaces

The last section of the dissertation, chapter VII, is concerned with problems in detecting and describing the surfaces that have been found to be so important. Two particular areas are addressed with further study suggested in certain areas. In both areas I am concerned with how noise affects the detection of surfaces from stereopsis. In the absence of r ise the task of detecting surface regions becomes much simpler since the surface can be found by looking for the absence of disparity contrast. With noise, however, there can be contrasting disparities at any location so some measure of the strength of points within a range of disparities must be used to know if a surface exists. This strength may be an absolute measure. That is, with a certain density of points the surface should be apparent independent of the amount of noise. Another possibility is suggested by the companion processes of facilitation and inhibition. With the combination of these processes the increase in strength of the surface is greater than linear. This suggests that a denser surface should have more resistance to noise than a sparse surface. The first experiment shows that this is the case. In this experiment, a random dot stereogram consisting of a planar surface parallel to the image plane is embedded in a certain percentage of points at random disparities. Subjects judged whether a surface was present in the image. The higher density surfaces were shown to be salient with a higher percentage of noise than the less dense surface.

Another factor which affects the detectability of surfaces is the type of surface. That is, properties of the surface affect the detectability of the surface just as they affect the way depth is perceived from the surface. For example, surface edge information may be useful in detecting the presence of a surface. The ability to resist noise is a measure of the strength of particular surface being tested. The second experiment used this property to compare the salience of different surface types by comparing their resistance to noise.

References

Barlow, H.B., Blakemore, C. & Pettigrew, J.D. (1967) The neural mechanism of binocular stereopsis. *J. Physiol., Lond.* **193**, 327-342.

Brookes, A. and Stevens, K.A. (1988a) The analogy between stereo depth and brightness contrast. In preparation.

Brookes, A. and Stevens, K.A. (1988b) Binocular depth from surfaces vs. volumes. In preparation.

Braunstein, M.L., Andersen, B.J., Rouse M.W, and Tittle J.S. (1986) *Percept. Psychophys.* **40**, 216-224.

Dosher, B.A., Sperling, G. and Wurst, S.A. (1986) Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Res.* **26**, 973-990.

Epstein W. (1973) The precess of 'taking-into-account' in visual perception *Perception* **2**, 267-285.

Foley, J.M. (1980) Binocular distance perception. *Psychological Review* **87**, 411-434.

Foley, J.M. and Richards, W. (1972) Effects of voluntary eye movement and convergence on the binocular appreciation of depth. *Percept. Psychophys.* **11**, 423-427.

Gillam, B J. (1968) Perception of slant when perspective and stereopsis conflict: experiments with aniseikonic lenses. *J. exp. Psychol.* **78**, 299-305.

Gillam, B.J., Flagg, T. and Finlay, D. (1984) Evidence for disparity change as the primary stimulus for stereoscopic processing. *Percept. Psychophys.* **36**, 559-564.

Gogel, W.C. (1956) Relative visual direction as a factor in relative distance perceptions. *Psychological Monographs* **418**, 1-19.

Gogel. W.C. (1972) Depth adjacency and cue effectiveness. *J. exp. Psychol.* **92**, 176-181.

Gulick, W.L. and Lawson, R.B. (1976) *Human stereopsis: A psychological analysis.* New York: Oxford University Press.

Hubel, D.H. & Wiesel, T.N. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (London)//fB 160,* 106-154.

Julesz, B. (1978) Global stereopsis: cooperative phenomena in stereoscopic depth perception. In *Handbook of Sensory Physiology VIII*, R. Held, H. Leibowitz, H. Teuber, eds. New York: Springer-Verlag.

Marr, D. (1982) *Vision.* San Francisco: Freeman.

Mayhew, J. and Longuet-Higgins, H.C. (1982) A computational model of binocular depth perception. *Nature* **297**, 376-379.

Mitchison, G.J. and McKee, S.P. (1985) Interpolation in stereo matching. *Nature* **315, 30**, 402-404.

Mitchison, G.J. and Westheimer, G. (1984) The perception of depth in simple figures. *Vision Res.* **24**, 1063-1073.

Morrison, J.D. and Whiteside, T.C.D. (1984) Binocular cues in the perception of distance of a point source of light. *Perception* **13**, 555-566.

Nelson J. (1975) Globality and stereoscopic fusion in binocular vision. *J. theor. Biol.* **49**, 1-88.

Nelson J. (1977) The placticity of correspondence: after-effects, illusions and horopter shifts in depth perception. *J. theor. Biol.* **66**, 203-266.

Ogle K.N. (1946) The binocular depth contrast phenomenon. *Am. J. Psychol.* **59**, 111-126.

Pastore N. (1964) Induction of a stereoscopic depth effect. *Science* **144**, 888.

Pastore N. and Terwilliger M. (1966) Induction of stereoscopic depth effects. *Br. J. Psychol.* **57**, 201-202.

Richards, W. (1977) Stereopsis with and without monocular contours. *Vision Res.* **17**, 967-969.

Ritter, M. (1977) Effect of disparity and viewing distance on perceived depth. *Percept. Psychophys.* **22**, 400-407.

Ritter, M. (1979) Perception of depth: Processing of simple positional disparity as a function of viewing distance. *Percept. Psychophys.* **25**, 209-214.

Rogers, B.J. and Graham, M.E. (1984) Aftereffects from Motion parallax and stereoscopic depth: similarities and Interactions In *Sensory Experience, Adaptation and Perception*, L. Spilmann and B. Wooten, eds. Hillsdale, N.J.: Lawrence Erlbaum.

Stevens, K.A. and Brookes, A. (1987a) Depth reconstruction in stereopsis. *Proceedings. First IEEE Conference on Computer Vision*, London, June.

Stevens, K.A. and Brookes A. (1987b) Probing depth in monocular images. *Biol. Cybernet.*. In press.

Stevens, K.A. and Brookes A. (1988) Integrating stereopsis with monocular interpretations of planar surfaces. *Vision Res.* **28**, 371-386.

Wallach, H. and Zuckerman, C. (1963) The constancy of stereoscopic depth. *Am. J. Psychol.* **76**, 404-412.

Werner, H. (1938) Binocular depth contrast and the conditions of the binocular field. *Am. J. Psychol.* **51**, 489-497.

Werner, H. (1942) Binocular Vision, normal and abnormal. *Archs. Ophthal., New York* **28**, 834-844.

Wheatstone C. (1838) Contributions to the psychology of vision, I: On some remarkable. and hithherto unobserved, phenomena of binocular vision. *Phil. Trans. R. Soc. Pt. 1* 371-394.

Vision Res. Vol. 28, No. 3, pp. 371-386, 1988
Printed in Great Britain

0042-6989/88 $3.00 + 0.00
Pergamon Press plc

INTEGRATING STEREOPSIS WITH MONOCULAR INTERPRETATIONS OF PLANAR SURFACES*

KENT A. STEVENS and ALLEN BROOKES
Department of Computer Science, University of Oregon, Eugene, OR 97403, U.S.A.

(Received 10 November 1986; in final revised form 11 September 1987)

Abstract—Experiments are reported that involved spatial judgments of planar surfaces that had contradictory stereo and monocular information. Tasks included comparing the relative depths of two points on the depicted surface and judging the surface's apparent spatial orientation. It was found that for planar surfaces the 3D perception was dominated by the monocular interpretation, despite the strongly contradictory stereo information. We propose that stereo information is effectively integrated only where the surface exhibits curvature features or edge discontinuities, i.e. where the second spatial derivatives of disparity are nonzero. Planar surfaces induce constant gradients of disparity and are thus effectively featureless to stereopsis. Further observations are reported regarding nonplanar surfaces, where contradictory monocular information can still be effectively rivalrous with that suggested stereoscopically.

Stereopsis Binocular vision Depth perception

INTRODUCTION

How does stereopsis constrain the perceived 3D shape and spatial orientation of static surfaces? The most plausible answer, seemingly, would be in terms of distance information determined from disparity at points across the surface. Stereopsis is generally expected to provide 3D distance information, specifically range and relative depth across visible surfaces, as derived from horizontal (and possibly vertical) retinal disparities given geometric parameters such as the angles of gaze and convergence (Mayhew, 1982; Longuet-Higgins, 1982a, b; Prazdny, 1983). There is much psychophysical evidence to support the view that stereopsis provides distance information. Stereopsis allows accurate judgments of absolute distance out to at least 2m (e.g. Wallach and Zuckerman, 1963; Ritter, 1977, 1979; Morrison and Whiteside, 1984), and, within that range, distance intervals are accurately perceived from disparity intervals (so-called "stereo depth constancy", see Ono and Comerford, 1977; Wallach *et al.*, 1979). It therefore seems reasonable to conclude that binocular vision in natural circumstances results in more-or-less complete and accurate 3D mapping of the surfaces in the immediate surrounds. But it is not clear how that 3D information might be combined with that derived monocularly.

Compared to stereopsis, the monocular "depth cues" in a static image provide much weaker and less precise 3D information†. Strongly restrictive assumptions are required to interpret cues such as shading, texture gradients, and monocular configurations such as in Fig. 1 (Stevens, 1981a, b, 1984). In comparison to the sound geometrical basis for determining absolute and relative distances from stereo disparity, one would expect stereopsis to dominate over the less reliable monocular information. This study and others, however, suggest the contrary: monocular configurations often dominate the resulting 3D interpretation over stereopsis, even in the near range where stereopsis is most accurate.

To be sure, binocular vision generally yields more accurate 3D judgments than monocular vision based on linear perspective, texture, shading, and so forth (e.g. Smith and Smith, 1957, 1961; Smith, 1965). Contradictory results were reported by Youngs (1976), however, where

*Supported by Office of Naval Research Contract N00014-K-84-0533.

†Monocular depth cues, despite their name, are primarily sources of information about local surface orientation (the orientation of surface patches relative to the line of sight) and of shape (surface curvature as well as the intrinsic geometry of the surface) and only in a weaker sense able to deliver distance information, either relative or absolute (Marr, 1982; Stevens, 1983b). That is, monocularly there is more reliable information about surface shape features and orientation than of distance *per se*.

Fig. 1. Monocular configurations that evoke definite 3D interpretations.

stereo disparity had no significant effect on apparent slant (of planar stimuli). Youngs (1976) questioned "why the disparity coding fails so miserably" in those experiments. Stereopsis is particularly weak in the presence of a strong contradictory monocular interpretation, such as presented in reversed-disparity stereograms of a face or a street scene (Wheatstone, 1852; Schriever, 1925; Gregory, 1970; Yellott and Kaiwi, 1979), or by Hochberg's striking Necker cube stereogram (see Julesz, 1971, p. 163), wherein a cube at constant retinal disparity readily reverses in depth.

We performed a series of experiments to attempt to determine what role stereopsis plays in the presence of contradictory monocular information. Experiment 1 concerned whether stereopsis could be used to effectively contradict the monocular interpretation of oblique intersections as foreshortened right angles, when the intersections were actually not perpendicular in 3D. We used stimuli similar to the planar grid in Fig. 1, and found stereopsis remarkably impotent in influencing the perceived orientation and 3D configuration. Experiment 2 similarly examined relative depth judgements in displays with conflicting stereo and monocular information. Given a simple pair of stereo points, that with the greater (more positive) disparity is seen as relatively farther. But if these points are embedded in a continuous 3D surface, and if the monocular interpretation suggests an alternative relative depth between the two points, that monocular interpretation governed the judgement in our experiment. Experiment 3 similarly examined whether a conflicting disparity gradient influenced the monocularly interpreted surface orientation.

We recognized a common theme: our stimuli, although rich in terms of stereo information, consisted of planar surfaces in 3D. Examination of control stimuli convinced us that sufficient stereo information was available, rather it appeared that stereo disparities across a planar surface were simply not effectively analyzed in 3D. More formally, we hypothesized that stereopsis extracts 3D surface information only where the second spatial derivatives of disparity are nonzero, corresponding to loci where the surface is curved, creased, or discontinuous. Experiment 4 directly examined planar versus nonplanar stereo stimuli, with and without competing monocular interpretations. The results further support this hypothesis. (And reviewing earlier studies, we observed that where stereopsis was particularly ineffective against conflicting monocular information, those studies involved planar surfaces.)

An adequate explanation must address two issues: the computation of depth from disparity and the integration of stereo and monocular 3D information. We will argue that depth is derived from disparity only where the surface exhibits continuous curvature or sharp discontinuities. But we suggest that depth, the apparent variation in surface relief, is reconstructed from multiple sources of evidence about surface topography. That is, surface shape is first analyzed in terms of sharp edges and creases, smooth folds, indentations, and so forth, from both binocular and monocular sources. The depth one experiences is a consequence of how this information is interpreted and reconciled. Depending on how the monocular information is interpreted, radically different depth distributions might be experienced. This is quite distinct from the notion that depth (and slant) is derived directly from stereo disparity (and its gradient).

EXPERIMENTS

Experiment 1: Interpretation of Perpendicular Intersections

Observers tend to interpret monocular images of oblique intersections as right-angle intersections in 3D (Attneave and Frost, 1969; Perkins, 1972; Shepard, 1981; Stevens, 1983a). In an earlier experiment, Stevens (1983a) found that subjects perceive such stimuli (e.g. a cross or a parallelogram) as lying on a plane oriented in 3D. Subjects could reliably visualize the orientation of that plane, and judge whether a line segment, superimposed on the monocular stimulus at a given image orientation, corresponded to the visualized normal to the plane. Moreover, apparent tilt (direction of slant) agreed closely with that predicted by assuming that the stimulus image corresponded to a right angle in 3D. In the present experiment we used similar cross and grid stimuli, but now projected stereoscopically, in order to examine whether the available stereo information would permit observers to distinguish the true 3D configuration.

Method

Apparatus. Stereo pairs were presented by a Wheatstone-style stereoscope using a pair of optically flat front-surfaced mirrors and two Tektronix 634 monochrome displays (flat 9 × 12 cm screens, 1100 line resolution, and less than 0.5% geometric distortion). The optic path from monitor screen to observer was 38 cm, and the two paths converged at total angle of 9.8° (providing consistent accommodation and vergence for a 65 mm interpupillary separation). Circular apertures allowed a 6.4° radius field of view. The stimuli consisted of luminous lines against a dark background. The stereograms were generated dynamically by a Symbolics 3670 Lisp Machine; the monochrome monitors projecting the left and right images were driven independently by separate channels of a color frame buffer.

To generate a stereo pair, 2D projections were computed from left and right vantage points that differed by the 9.8° convergence angle. The images could be generated in either perspective or orthographic projection. In the perspective case (used in Experiments 2 and 3) the projection was computed as if the surface were physically situated 38 cm from the viewer; for the orthographic case (Experiments 1 and 4) the viewing distance was 100-fold further with the image scaled accordingly so as to subtend the same visual angle as in the perspective case. All computed stereo disparities were distributed equally to the two half-images, corresponding with a frontal, foveal viewpoint with symmetrical convergence of the two eyes.

Stimuli. Two types of orthographic stimuli were presented stereoscopically: a pair of crossing lines and a 5 × 5 grid of lines. The angle of intersection was either 90° (Fig. 2) or skewed 15, 30 or 45° from the perpendicular (Fig. 3). The grid became an increasingly racked parallelogram with increasing skew angle. Monocularly, varying skew angle would imply different spatial orientations; stereoscopically the spatial orientation should remain constant and only the intersection angle should appear to vary. The intention was to place a compelling monocular* impression of perpendicularity in opposition to contradictory stereo information. Note that orthographic projection was used to avoid a monocular cue to skew angle provided by perspective distortion to the skewed grid.

The stimuli were specified by three spatial parameters relative to the plane containing the grid or cross. The orientation of the plane in stereo was defined by its slant (the angle between the normal to the plane and the line of sight) and tilt (the direction to which the normal would project, i.e. the direction of slant). The third parameter specified the angular orientation of the grid or cross on the slanted plane (a rotation about the normal to the plane). The slant was held constant at 65°. Three angles of tilt and two angular orientations were used to provide six visually distinct perspectives of the grid and cross stimuli for each of the four skew angles—see (Stevens, 1983a) for similar cross and grid experiments in which the accuracy of apparent tilt judgments was found to be substantially independent of the choice of tilt angle.

Procedure

Ten graduate students participated as paid subjects; all had good stereo vision and were naive to the purposes of the experiment. The subjects were shown example stimuli and explained that they would see crosses and grids oriented at a slant relative to the observer and that the 3D intersections would sometimes be right angles and at other times skewed (the notion of a skewed intersection was reinforced with a physical demonstration). They were to make force-choice judgments of whether the intersection was perpendicular in 3D or not (referred to as the P judgment, made by depress-

*Here we refer to the fused binocular image as a 2D projection, in Julesz's (1971) sense of a "cyclopean" retina. The projection might be described geometrically as the average of the left and right half images, or the equivalent projection that would arise with a zero interpupillary separation. We will refer to the "monocular" information present in that projection, disregarding the disparity information that is present as well.

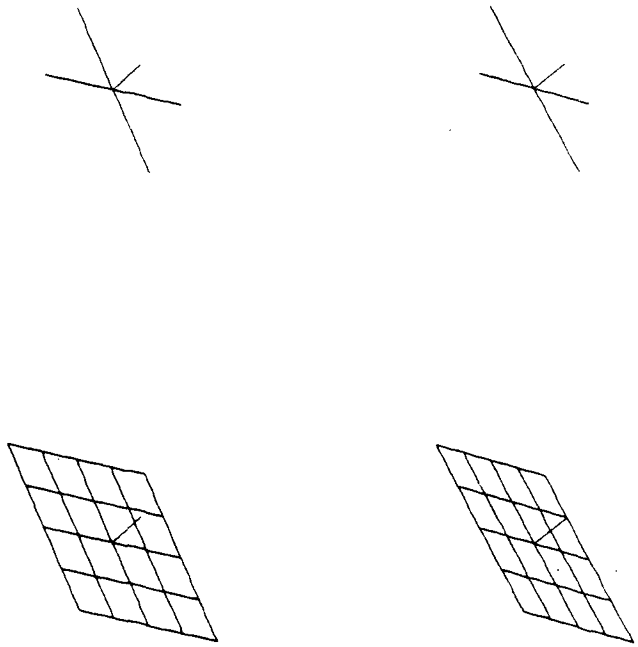


Fig. 2. Examples of cross and grid stereograms, each with 0° skew angles. Note that the normal appears to project perpendicularly to the plane defined by the cross or grid.

ing a mouse button). A positive response corresponded to lines that appeared within approximately 5° of perpendicular. Unlimited presentation time was allowed. The P judgment response initiated the addition of a stereo line segment to the stimulus that was a geometrically accurate rendition of the normal to the plane of the cross or grid. The subject made a second

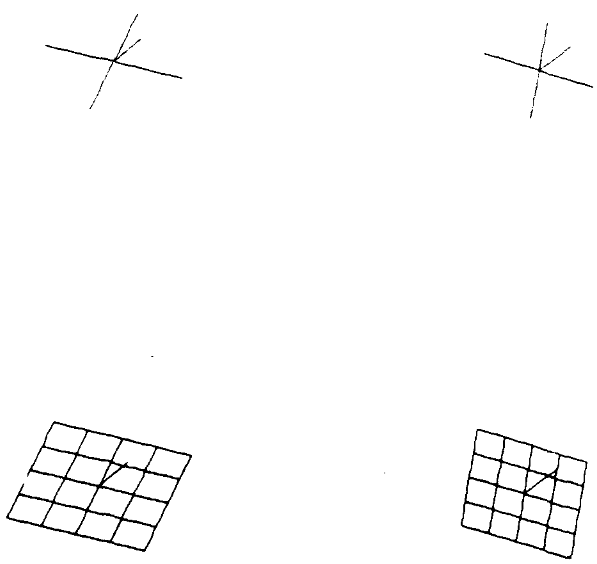
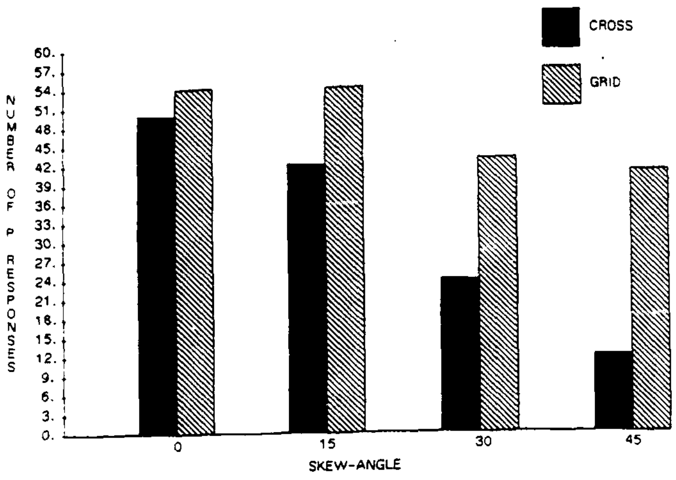
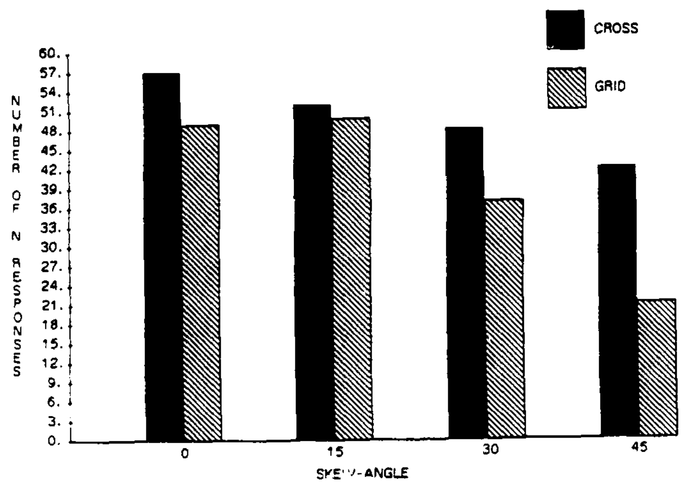


Fig. 3. Cross and grid stereograms, with identical spatial orientation as in Fig. 2, but with intersections skewed 45° from perpendicular. Note that the "normals" do not appear perpendicular to the plane of the grid or cross.



A



B

Fig. 4. Judgments of perpendicularity as a function of skew angle for cross and grid stimuli in (a); corresponding judgments of the surface normal in (b).

forced-choice response whether the line appeared to be normal (the N judgment, with the same criterion of roughly 5°).

Results and discussion

Figures 4(a) and (b) graph the number of P and N judgments as a function of skew angle for the cross and grid stimuli. For 0° skew the monocular and stereo information are both consistent with right angle intersections on a plane slanted 65°. Hence the 0° skew condition provides a baseline for the P and N judgments at greater skew. As skew angle increased, the N and P judgments for crosses and grids showed

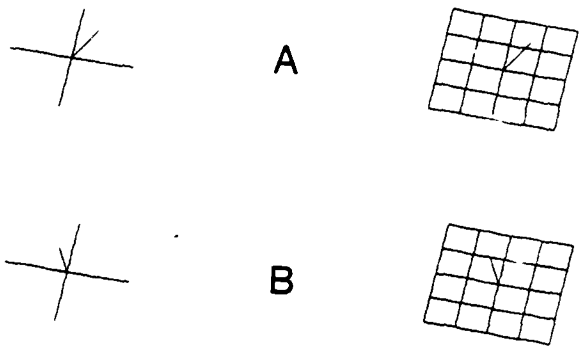


Fig. 5. In (a) the normal is correct for the monocular projection of a cross skewed 45°. In (b) the normal is correct of the monocular projection of a right angle intersection.

different, and complementary, trends. Concerning the P judgments, the grids had a greater tendency to be seen as perpendicular, and correspondingly, the displayed normals appeared increasingly incorrect as skew angle increased. The crosses were seen more vertically (i.e. according to the stereo information) although both P and N decreased with increasing skew for the crosses as well. Overall the grids were much more persistently judged on the basis of the monocular information. These trends all showed significance at $P < 0.05$ using sign tests comparing the N and P judgments for 0° and 45° skew angles.

Since the stereo projection of the normal was geometrically correct with regard to the plane containing the intersecting lines, regardless of their angle of intersection in 3D, if stereopsis had dominated the P and N judgments, the intersections would have appeared skewed for all but the 90° case and the normals would have always appeared correct. Conversely, if the judgments were based on the monocular information, the intersections would have always appeared perpendicular and the normal would have appeared incorrect except for the 90° case. The data fell between these two alternatives: the monocular interpretation was markedly influential despite the geometrically-correct stereo information, and significantly more so for the grid than the cross. We also note that the subjects' overall ability to judge the intersection angle was not particularly sensitive (e.g. skew angles differing by 15° were barely distinguishable).* Thus the lack of precise correspondence between the N and P judgments as a function of skew angle may reflect the differences in difficulty of the two tasks.

Figure 5(a) depicts the tilt of the surface normal for a cross and grid that is skewed 45°. This figure was rendered by projecting an experimental stimulus, with the geometrically-correct surface normal, at 0° rather than 9° convergence angle. Note that the normal in Fig. 5(a) seems incorrect. Figure 5(b), which appears more appropriate, was computed by assuming the projection corresponds to a square cross or grid (see Stevens, 1983a, appendix, for formula). Figure 5(b) thus illustrates the difference between the geometrically-correct stereo interpretation of a 45° intersection, and what one would perceive if that intersection were assumed perpendicular.†

Given the richer stereo information in the grid stimulus (10 lines and 25 intersection points, compared to 2 lines and one intersection point) one might expect more accurate spatial localization of the grid than the cross. But stereopsis had a weaker role in determining both the perceived 3D orientation of the grid and the angle of intersection of the grid lines, compared to the simpler cross stimulus. There was seemingly a greater tendency to "ignore" the stereo information in the grid compared to the cross stimuli.

*We later asked two experienced observers to judge the angle of intersection for various cross stimuli and found that they could accurately estimate the true intersection angle to within 5° or so, and yet, for the correspondence grid stimuli, they repeatedly judged a 45° intersection to be skewed only 15° or so from perpendicular.

†Quantitatively, the difference in tilt amounts to 64°. The slant is also influenced by assuming the intersection is 90°. For example, the grid stereogram in Fig. 3 appears slanted much less than 65°). The computed monocular slant for Fig. 3, assuming it corresponds to a square grid, is only 38.5°.

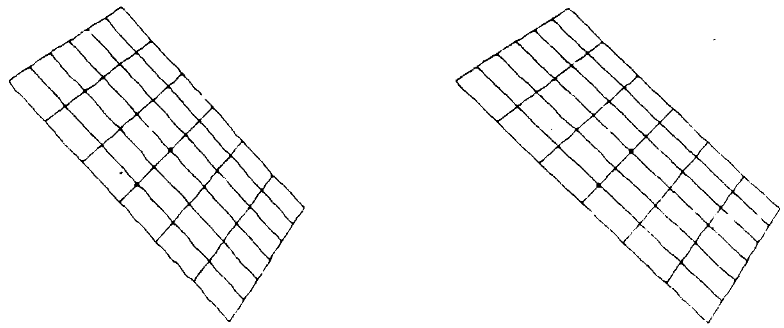


Fig. 6. Example stimulus in which subjects judged whether the given probe point was nearer than, equidistant, or further than the central reference point. The stereo disparity gradient was either consistent with, orthogonal to, or opposite from the monocularly implied distance gradient.

Experiment 2: Two-Point Relative Depth Judgments

Method

Stimuli. The optical arrangement was unchanged from Experiment 1, but we now decoupled the computation of stereo disparities from the monocular projection of the individual half-images. The aim was to examine the influence of conflicting stereo and monocular information on the judgement of the relative depth of two points on the depicted surface. The stimulus surface was a 7 × 7 square grid of lines projected in perspective, slanted 65° as in Experiment 1, and tilted either 45 or 135°.

To compute the stereogram, the screen coordinates of the two half-images were first projected according to a 0° vergence angle, which would have resulted in identical half-images, except for the introduction of horizontal disparities that were either consistent or inconsistent with the monocular projections. Four cardinal directions were defined on the stimulus surface, with north corresponding to the monocular direction of tilt (i.e. distance increased to the north on the basis of perspective). The stereo and monocular information were consistent when the stereo disparity gradient was northward. When the gradient increased to either the east or west it was orthogonal to the monocular perspective, and when to the south the stereogram had effectively reversed disparities. The surface at the central reference point always had zero disparity.

Procedure. The four subjects had participated earlier in the first experiment. The task was to judge whether a given probe point was nearer or further than, or at the same depth as a reference point located at the center of the surface. The probe point was 6° away from the reference point in one of the four cardinal directions (Fig. 6). Both probe and reference points subtended 10′ and were projected stereoscopically with disparities corresponding to points embedded in the stereo surface of the grid. There were 5 repetitions of the 32 stimuli: 2 tilts (45° and 135°), 4 probe locations (N, S, E, W), and 4 directions for the disparity gradient, in random order.

Results and discussion

Table 1 shows the sets of relative depth responses for each combination of probe

Table 1. Percentage of judgments that the probe point appeared nearer than (<), equidistant (=), or farther than (>) the central reference point. The relative depth predicted on basis of stereo disparities is in bold

Direction of disparity gradient	Probe location N			S			E			W		
	<	=	>	<	=	>	<	=	>	<	=	>
North	0	0	**100**	**100**	0	0	25	**53**	22	18	**55**	27
South	**3**	12	85	92	8	**0**	8	**67**	25	22	**70**	8
East	0	**0**	100	100	**0**	0	33	42	**25**	**18**	60	22
West	0	**13**	87	87	**13**	0	**18**	60	22	22	63	**15**

Fig. 7. The disparity gradient is perpendicular to the apparent monocular gradient of distance. Subjects adjusted the monocular "normal" by rotating it in the image plane until it appeared perpendicular to the grid in 3D, i.e. to align with the surface normal.

location and disparity gradient direction. The values in boldface indicate the responses consistent with the stereo disparities. The first row serves as a control, since the direction of the stereo and monocular gradients coincided. For this case the N and S probe locations show the expected depth judgments. The E and W probe locations were generally judged equidistant, but there were also several "farther than" and "nearer than" judgments. The "equidistant" judgment turned out to be problematic. Since the two half-images were projected in perspective, points due east and west of the central reference point would have been necessarily farther than the reference point simply by the perspective projection. We thus carefully computed the E and W probe locations to be slightly south of due east and west so that, monocularly, they and the reference point were equidistant from the observer. Nonetheless it turned out rather difficult to decide whether the E, W, and reference points appeared equidistant, even with consistent stereo information, and even for highly experienced observers.

When disparity was reversed (Table 1, second row) there was an overwhelming tendency to continue to see the N point as farther, and the S point as nearer, that is, according to the monocular perspective and contrary to the stereo disparities. Some "regression to the frontal plane", is apparent, suggesting that subjects experienced a reduced impression of depth or slant in this case, as Gillam (1968) also found in reversed-disparity stereograms.

The important cases, we believe, concern disparity gradients *orthogonal* to the monocular distance gradient. Consider, for example, the case of the disparity gradient to the west and the probe point west of the reference point. The probe had positive disparity, and on that basis should have been seen as farther, but was not. The direction of the disparity gradient had no systematic effect on the depth judgments for the east and west probe locations. Overall, the apparent depth corresponded very closely with the monocular perspective, despite the contradictory stereo information.

Experiment 3: Surface Orientation Judgments

The results of Experiment 2 suggested that a disparity gradient orthogonal to the perspective distance gradient had negligible influence on the relative depths of two points on the surface. Experiment 3 pursued this result in terms of the effect of a competing disparity gradient on apparent tilt—see method in (Stevens, 1983a) Subjects adjusted a needle to appear perpendicular to the apparent plane of the grid. If the orthogonal disparity gradient had an effect, we would expect the needle to lean in the direction of the stereo gradient, an effect analogous to the vector sum of the monocular and stereo interpretations.

Method

Stimuli. Stereograms were constructed for which the stereo information corresponded to a surface whose 3D orientation was precisely orthogonal to that depicted monocularly. The stimulus surface was a 5 × 5 square grid of lines projected in perspective, slanted either 35 or 70° and tilted either 40 or 140°. The disparities corresponded to a slanted plane whose tilt was ±90° away from the monocular tilt. The monocular cue implied depth increasing to the north while the stereo information implied depth increasing to either the east or west, depending on the polarity of the disparity gradient.

Procedure. Three subjects were used; all had previous experience in the experimental series. A grid surface was presented for one second prior to superimposing a rotatable line segment that had one endpoint fixed at the center of the grid. The "needle" was presented monocularly, to the

Table 2. Mean surface tilt judgments (and standard deviations) with monocular normal

Slant	Tilt	Disparity gradient to west	Disparity gradient to east
35.0	40.0	50.5 (5.3)	49.2 (7.5)
35.0	140.0	142.7 (4.5)	141.2 (4.5)
70.0	40.0	46.8 (2.2)	44.0 (4.0)
70.0	140.0	139.8 (3.5)	138.7 (3.6)

dominant eye only (see Fig. 7). Subjects stepped the needle in tilt by $\pm 2.5^\circ$ increments until it pointed in the direction of the surface normal. The needle appeared to emerge from the surface and to pivot in 3D about the fixed end, despite only rotating in the image plane. Unlimited time was permitted per trial. Tilt data was recorded for 5 trials of each of 8 conditions (four monocular surface orientations times two directions for the stereo disparity gradients).

Apparent slant was also recorded using a stereoscopic needle that could be stepped in both slant and tilt. The three subjects were presented 5 trials per each of the eight conditions, as above.

Results and discussion

Since the disparity gradient was orthogonal to the monocular depth gradient, the apparent normal might be expected to lean in the direction of the disparity gradient, e.g. to rotate counterclockwise (increase numerically) when the disparity gradient was to the west, and clockwise when the gradient was reversed to the east. However, the data exhibited no systematic leaning in the direction of the stereo disparity gradient (see Table 2). Moreover, the apparent tilt was in reasonably close agreement with the monocularly predicted tilt. Overall, the apparent tilt seemed determined only monocularly.

Similarly, apparent slant was in close accordance with that predicted by the monocular perspective (see Table 3). This is remarkable given that the stereo disparity was constant (and zero) in that direction. The slant probe was adjusted to within one standard deviation of the slant suggested by the monocular perspective for all conditions.

Stereo disparity was constant in the direction that the monocular cues indicated increasing depth, and vice versa. With the two cues orthogonal, if they were somehow summated, one would expect the resulting apparent tilt to be influenced by the direction of the disparity gradient, but no such effect was observed. Moreover, apparent slant was in good correspondence with that predicted by the monocular perspective, despite the fact that stereo disparities were constant in that direction. This experiment thus extends the more qualitative findings of Experiments 1 and 2.

Experiment 4: Planar vs Nonplanar Stereo Disparity Distributions

In this final experiment we used line grid and random dot stereograms of planar and nonplanar surfaces to explore the importance of surface geometry on the simple two-point relative depth judgment (as in Experiment 2) in the presence and absence of competing monocular information. Our strategy was to embed a pair of stereo points in various surfaces to see to what extent the "context" influenced the apparent relative depths of these two points.

Method

Stimuli. The stimuli were grid stereograms (with lines separated by 1.9°) and random dot stereograms (Fig. 8). The horizontal disparity across the stereogram was a continuous one-dimensional function of screen position, corresponding to either a slanted plane, a Gaussian ridge, or a Gaussian-smoothed edge. These "stereo surfaces" were oriented either horizontally (h) or vertically (v). The slanted plane v, for example, corresponded to a plane pivoted about the vertical meridian, with disparities that varied from $0'$ at the center to $\pm 51.2'$ at left and right extremes of the field of view (occluded by the optical apparatus at 6.4° eccentricity). Similarly, the Gaussian ridge function induced stereo disparities from $-37.8'$ along the ridge to $0'$ in the periphery [see the horizontally oriented ridges in Fig. 8(a) and (b)]. The ridge protruded towards the viewer with half-amplitude at $\pm 1.6^\circ$ eccentricity. The Gaussian-smoothed edge had the same space constant as the ridge. It presented a smoothed step transition from $\pm 18.9'$ at opposite edges of the field that passed through zero along the vertical or horizontal meridian [see vertical case in Fig. 8(c) and (d)].

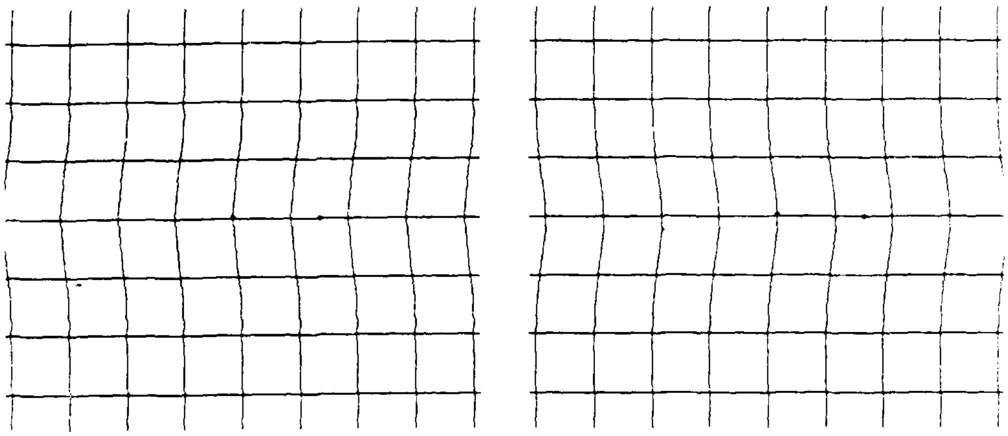
Table 3. Mean surface slant judgments (and standard deviations) with stereoscopic normal

Slant	Tilt	Disparity gradient to west	Disparity gradient to east
35.0	40.0	36.5 (2.8)	37.5 (4.2)
35.0	140.0	33.0 (3.6)	38.8 (6.1)
70.0	40.0	68.5 (4.6)	64.7 (8.7)
70.0	140.0	65.0 (7.7)	66.5 (6.6)

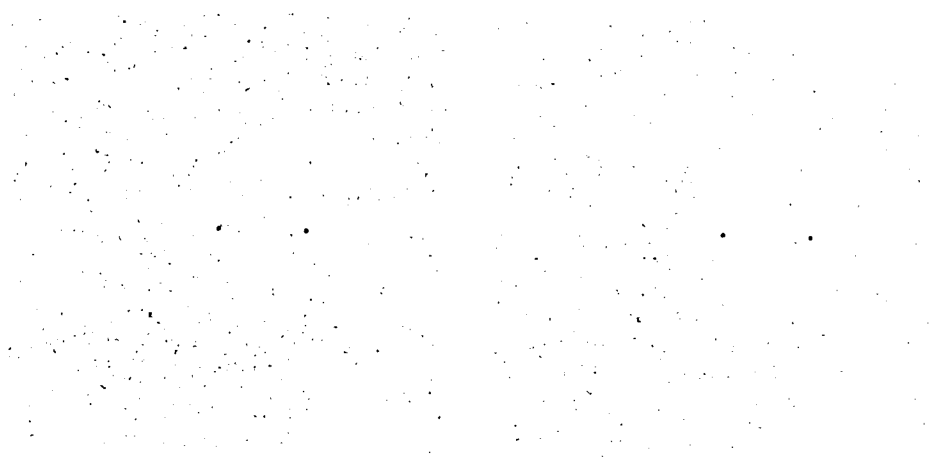
Procedure. Three subjects from earlier experiments were used; all had excellent stereo vision. The task, as in Experiment 2, was to judge the depth of a probe point relative to a central reference point. The probe and reference points both subtended 10′. The probe was placed at 2.9° eccentricity either north (above), south, east (right of), or west of the reference point. The probe and reference points were both on the given stereo surface. (For the Gaussian ridge *h*, for example, the reference point had −37.8′ disparity. The probe point had 0′ disparity when north or south and −37.9′ when east or west of the reference point.) The subject indicated by mouse button whether the probe point appeared nearer, at the same depth as, or farther than the reference point. Free eye movements and unlimited observation time were allowed. The grid and dot versions of the experiment were run separately, each with 5 repetitions of the 24 conditions (six oriented disparity surfaces times four probe locations) in random order.

Results and Discussion

The relative depths of two stereo points could be determined, in principle, directly from their corresponding disparities. In a pilot experiment, where only the probe and reference points were displayed against a black background, their relative depth could be judged immediately and accurately, in accordance with their relative disparities. But when the two stereo points were embedded in a stereo surface, we found that the depth judgment depended on that surface. We conjecture that the depth judgment was mediated not directly by the relative disparities but by the perceived depth of the underlying surface. And, the perceived depth of the surface is

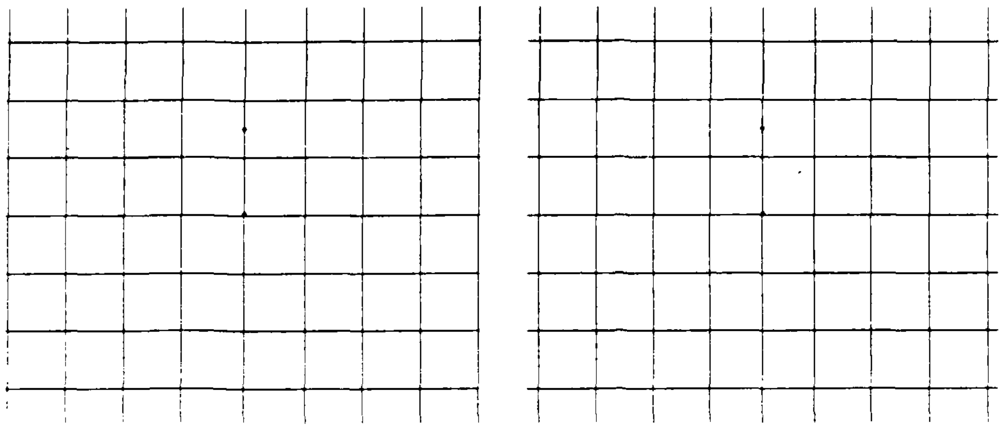


A

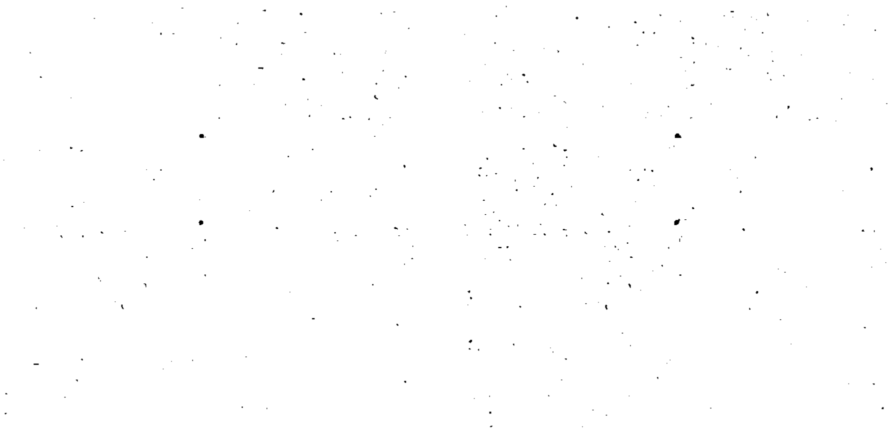


B

Fig. 8(A,B)



C



D

Fig 8. Horizontal Gaussian ridges in (A) and (B), vertical Gaussian edges in (C) and (D)

not strictly determined by the disparity distribution.

Table 4 shows the responses for the grid stimuli. The values in boldface indicate depth judgments consistent with the relative stereo disparities. Consider the case of the slanted plane *h*, where disparity increased from south to north. The N probe location should have been seen as farther, but zero "farther than" judgments were in fact recorded, and likewise zero "nearer than" judgments for the corresponding S probe location.* Similar results were obtained for slanted plane *v*. It is remarkable that when the dots were embedded in a surface which had a constant gradient of disparity the apparent relative depth of the probe and reference dots collapsed. Points that were readily seen as lying at different depths when viewed in isolation appeared equidistant when embedded in the constant gradient, but seemingly unslanted, grid.

For the nonplanar cases, the edge and ridge, the data are in better accordance with the stereo information, and generally better for the *h* than the *v* surfaces. This anisotropy has been reported earlier by Tyler (1973), Wallach and

*Several of the relative depth responses were actually *opposite* that predicted by the stereo disparities. We conjecture that this was due to illusory linear perspective caused by stereo depth constancy compensation. While the grid lines were horizontal and vertical in each half-image, the fused grid appeared to be trapezoidal rather than rectangular, presumably because of apparent length was scaled with increasing disparity. The rectangular grid appeared distorted by linear perspective. The slant implied by the perspective, of course, was opposite that implied by the disparity gradient. This effect suggests to us that stereo size constancy operates independently of processes responsible for apparent depth.

Table 4. Percentage of responses that the probe point is nearer than (<), or equidistant (=), or farther than (>) the central reference point, as in Table 1. The probe and reference points were both embedded in a stereo surface, in this case rendered by a square grid (see Fig. 8). The relative depth judgment predicted by the relative stereo disparities is in bold

Stereo surface and orientation	Probe location N			S			E			W		
	<	=	>	<	=	>	<	=	>	<	=	>
Slanted plane *h*	33	67	**0**	**0**	73	27	0	**100**	0	0	**100**	0
Gaussian edge *h*	0	0	**100**	**93**	0	7	0	**100**	0	0	**100**	0
Gaussian ridge *h*	0	0	**100**	0	0	**100**	0	**100**	0	0	**100**	0
Slanted plane *v*	27	**73**	0	0	**93**	7	0	80	**20**	**0**	100	0
Gaussian edge *v*	20	**80**	0	0	**87**	13	0	60	**40**	40	53	7
Gaussian ridge *v*	20	**80**	0	0	**93**	7	0	60	**40**	0	87	**13**

Bacon (1976), and Gillam *et al.* (1984) in depth detection tasks and by Rogers and Graham (1983) in the Craik–O'Brien–Cornsweet effect for stereopsis. Note that the depth of the Gaussian edge *v* was detected with only slightly better success than the slanted plane *v*.

We conclude that while depth can be encoded "directly" from disparity for isolated disparity points, *when those point are perceived as lying on a surface, their depth depends on the perceived depth of the surface*, which might happen to be negligible, either because it is a "featureless" field of stereo points in the absence of monocular 3D cues, or there are contradictory monocular cues.

The dramatic influence of the monocular grid is apparent in comparing the grid data in Table 4 with the corresponding random dot surface data in Table 5. The grid seemingly masked or "flattened" the depth undulation indicated by the disparity values. Significantly, the depth in the slanted plane stimuli, particularly in the *v* orientation, remained more difficult to detect than in the ridge and edge stimuli, even in the absence of a contradictory monocular 3D interpretation (of an unslanted rectangular grid). Ninio and Mizraji (1985) similarly observed that structured stereograms are less accurately perceived in 3D than unstructured (they used rectilinear grids as well). We interpret this as due to the conflicting monocular interpretation provided by the grids beyond the issue of the ineffectiveness of planar disparities.

GENERAL DISCUSSION

The 3D interpretation in these binocular stimuli was governed largely by the monocular cues. This is not to be construed as evidence of simple dominance of monocular over stereo cues, however. Instead, we believe that these planar stimuli happened to be particularly rich in monocular 3D cues, especially perspective and foreshortening, and particularly poor in stereo information due to our relative insensitivity to constant disparity gradients in the absence of disparity contrast. Stereo depth derives most effectively from disparity contrast: when disparity varies linearly it is dramatically less salient, despite large overall variations in disparity. In the absence of competing monocular cues a uniform gradient of disparity does effectively yield stereo depth, thus we do not conclude that stereopsis is wholly "blind" to constant disparity gradients. Rather, we suggest that depth interpretation from stereopsis is effectively reconciled with that from other sources primarily in terms of surface curvature and depth discontinuity features, and since our stimuli were devoid of these features, the monocular interpretation dominated.

Table 5. Relative depth judgments, as in Table 4, but for a surface depicted by a dense random dot pattern (see Fig. 8)

Stereo surface and orientation	Probe location N			S			E			W		
	<	=	>	<	=	>	<	=	>	<	=	>
Slanted plane *h*	27	0	**73**	**73**	7	20	0	**73**	27	0	**67**	33
Gaussian edge *h*	7	13	**80**	**100**	0	0	0	**100**	0	0	**93**	7
Gaussian ridge *h*	0	0	**100**	0	0	**100**	0	**100**	0	0	**100**	0
Slanted plane *v*	40	**60**	0	0	**60**	40	0	7	**93**	**53**	7	40
Gaussian edge *v*	0	**100**	0	0	**100**	0	0	0	**100**	**100**	0	0
Gaussian ridge *v*	0	**100**	0	7	**93**	0	0	0	**100**	0	0	**100**

The fact that stereo depth must compete with monocular depth even in simple experimental stimuli likely accounts for several depth phenomena reported earlier. Westheimer (1979) and McKee (1983) observed that when two vertical lines, projected at different disparities, are connected by horizontal lines to form a square, the threshold for detection of the depth difference is greater than when only the two vertical lines are presented. McKee (1983) suggested that the effect was due to the lines being connected into a perceptual whole. Mitchison and Westheimer (1984), studying variations on this configuration, demonstrated that the detection thresholds were elevated most when the disparities varied linearly (according to a slanted plane). They use the term "salience" to refer to a local weighted sum of disparity first differences between a given point and its neighbors which scales roughly inversely with the separation of stereo features. [This notion quantifies Gogel and Mershon's (1977) "adjacency effect".] Accordingly, local variations in salience (i.e. second differences of disparity) would reveal deviations from planarity in the corresponding surface. A slanted plane would present points of equal salience, and consequently of zero apparent variation in depth. Gillam *et al.* (1984) observed, in these terms, that depth derives most readily from places of high "salience".

But Mitchison and Westheimer (1984) also said that more is involved in the perception of depth from disparity, since their proposal cannot account for the dramatic extinction of depth in the simple case of the slanted square compared to only the vertical lines of the square. McKee (1983) regarded this as a figural connectivity issue, recall. We believe McKee was close to the mark: it is not the connectivity *per se* that is important (as Mitchison and Westheimer demonstrated) but the fact that the connectivity helped induce a monocular figure, a square, that has a compelling 3D interpretation. The square suggested a plane of zero slant, which dictated that the two vertical sides of the plane are equidistant from the viewer. The following illustrates the dramatic influence a monocular interpretation has on the eventual depth percept.

An ellipse, seen from a particular viewpoint, foreshortens to a circle in orthographic projection—e.g. an ellipse of 2:1 aspect ratio rotated about its minor axis to a slant of 60°, so that the major axis foreshortens by a factor of 0.5 (the cosine of 60°). A 2:1 rectangle would likewise foreshorten to a square. The stereograms in Fig. 9 depict concentric ellipses (and rectangles) lying on a plane of 60° slant. A compelling monocular 3D interpretation would be of a tunnel or funnel extending in depth from periphery to center. Seven subjects, naive to the experimental design, interpreted the stereograms accordingly, with the innermost circle (or square) seen as further than the outermost. While some observers noted that the outermost circle (or square) appeared slightly slanted, the apparent slant vanished towards the innermost. Apparent depth increased radially towards the center of the pattern rather than from right to left, despite the fact that the vertical meridian was at zero disparity. When the subjects were subsequently told that the stimuli corresponded to foreshortened ellipses and rectangles lying on a slanted plane, some subjects could see the slanted plane, while curiously others could not.

Figure 10 is, we believe, a particularly effective demonstration of the monocular influence. The lines are coplanar, i.e. increase linearly in disparity from left to right. The 3D impression, however, is of a corridor extending in depth, bordered on either side by columns of vertical lines or stakes. In the apparatus the innermost lines on either side of the vertical meridian had stereo disparities of $\pm 11'$; the outermost lines had disparities of $\pm 51'$. It is remarkable that the line with $-11'$ disparity appeared more distant than the line of disparity $+51'$. This apparent disregard for stereo disparity is far more blatant than that reported by Mitchison and Westheimer (1984), where thresholds were elevated by only a few minutes of arc. The difference, we suggest, is that figure 10 offers a far more compelling monocular 3D interpretation. But it is also noteworthy that experienced stereo observers can also discern the true stereo depth of the component lines with scrutiny, especially in Fig. 10, as if the monocular depth interpretation can be selectively disregarded.

The final observation we offer concerns interactions between stereopsis and monocular interpretations in the case where the stereo disparities suggest a highly salient curvature feature. In Fig. 11 the monocular interpretation is of a slanted plane, but the stereo disparities correspond to a 2D Gaussian in depth protruding towards the viewer. Note that the disparities are symmetrically distributed over the two half-images so that the fused "cyclopean"

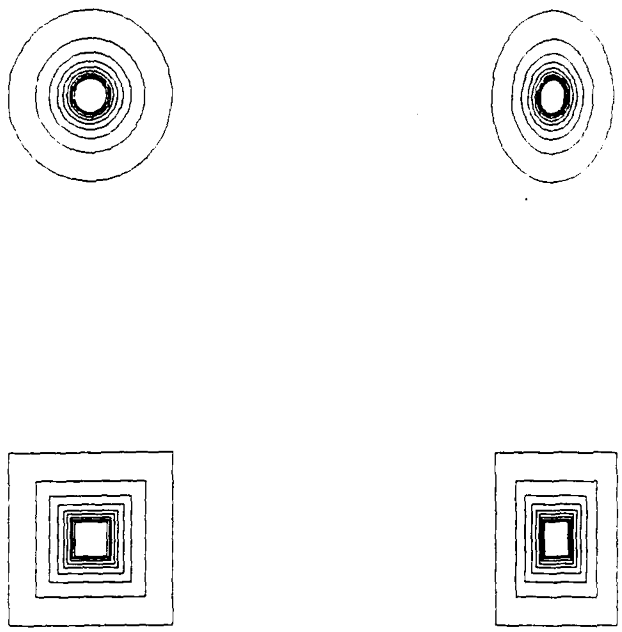


Fig. 9. Coplanar ellipses and rectangles, 2:1 aspect ratio and slanted 60°, in orthographic stereoscopic projection. A compelling monocular interpretation is of tunnels with circular and square cross-section seen in perspective.

Fig. 10. Lines on a common plane slanted 60°, but seen as a corridor in depth, as suggested monocularly.

image consists of straight lines, suggesting a slanted rectangular grid in perspective. We find that observers vary considerably in their interpretation of such a rivalrous figure, some seeing only a slanted plane, others seeing a plane at first then gradually becoming aware of a phantom protrusion in the center of the stereogram. Others achieve the nonplanar interpretation only after studying the random-dot stereogram version of the same Gaussian-shaped feature (Fig. 12) then re-examining the grid stereogram. Depth appears to be the end consequence of

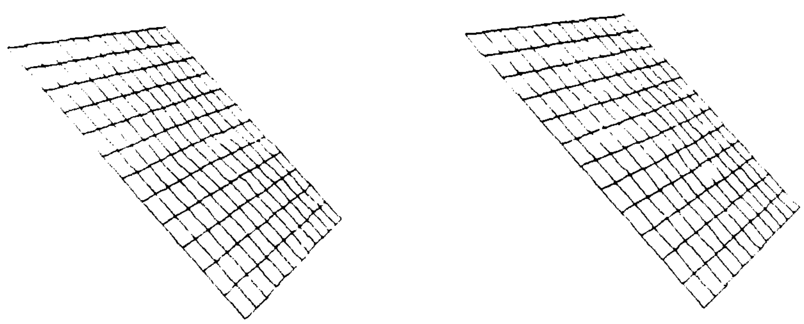


Fig. 11. A rivalrous pattern, monocularly a slanted plane, and stereoscopically a 2D Gaussian in depth.

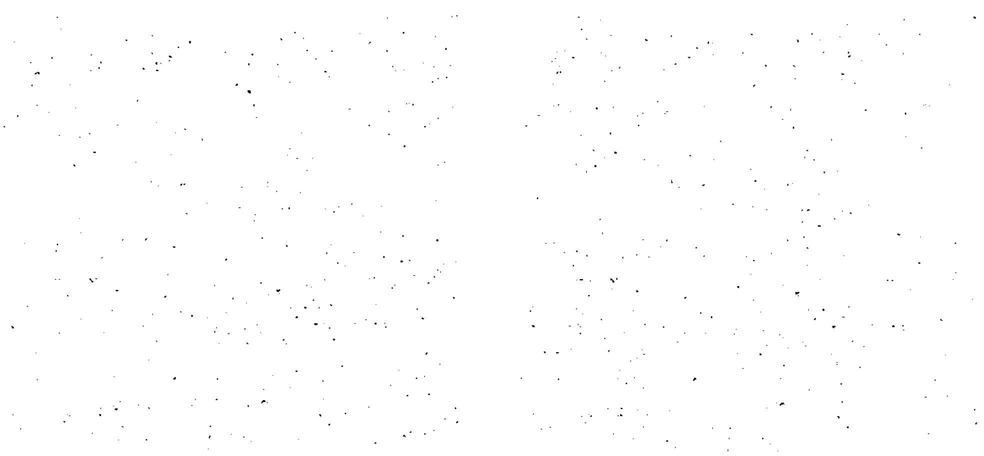


Fig. 12. The random-dot stereogram of the Gaussian in depth in Fig. 11.

a process that involves substantial "inference" or interpretation, that one sees depth according to the interpretation of 3D surface shape that one imposes. In that regard stereopsis is but one source of 3D shape information, and not necessarily the compelling one.

This series of experiments suggests that monocular cues have a stronger role in 3D perception than perhaps has been assumed. Likewise, stereopsis plays a much weaker role in the determination of depth across planar surfaces than expected. For very simple stereograms, an isolated pair of lines or points, say, the depth is indeed governed by the stereo disparities. But the contribution of stereopsis to the 3D percept changes dramatically as the stereogram is made more complex. With sufficient disparity evidence to suggest a continuous surface it is the *spatial distribution* of disparities, and not their individual magnitudes, that governs the apparent shape and depth. Specifically, the spatial distribution is analyzed to detect curvature and sharp discontinuities. Planar arrangements of disparity are in this regard featureless. This conclusion is close to that of (Gillam *et al.*, 1984) and (Mitchison and Westheimer, 1984) regarding the weak apparent depth associated with constant disparity gradients. In work reported elsewhere (Stevens and Brookes, 1988) we further conclude that surface curvature and discontinuity features are the primitive surface descriptors with which the visual system integrates stereo information with that contributed from monocular sources. In terms of spatial derivatives, we propose that the effective stereo features correspond to places where the second spatial derivatives are non-zero. The corollary is that neither the gradient (first spatial derivatives) nor the zeroeth derivatives (the raw disparity values themselves) are accessible as local surface shape descriptors. That is, neither slant nor relative depth is extracted directly from the disparity distribution across a surface. But, we must emphasize, relative depth is extracted from simple discontinuous configurations, such as between discrete, isolated items and across edges. And binocular vision undeniably provides absolute range information as well, particularly from convergence angle (Ritter, 1979) and in conjunction with motion parallax (Johansson, 1973). But we propose that range perception, which is most accurate in the near field (up to 2 m) under conditions of precise stereoscopic fixation, subserves motor functions such as locomotion and manipulation and *not* the perception of surface relief or 3D shape.

REFERENCES

Attneave F. and Frost R. (1969) The determination of perceived tridimensional orientation by minimum criteria. *Percept. Psychophys.* **6**, 391–396.

Gillam B. J. (1968) Perception of slant when perspective and stereopsis conflict: Experiments with aniseikonic lenses. *J. exp. Psychol.* **78**, 299–305.

Gillam B. J., Flagg T. and Finlay D. (1984) Evidence for disparity change as the primary stimulus for stereoscopic processing. *Percept. Psychophys.* **36**, 559–564.

Gogel W. C. and Mershon D. H. (1977) Local autonomy in visual space. *Scand. J. Psychol.* **18**, 237–250.

Gregory R. (1970) *The Intelligent Eye.* Weidenfeld & Nicolson, London.

Johansson G. (1973) Monocular movement parallax and near-space perception. *Perception* **2**, 135–146.

Julesz B. (1971) *Foundations of Cyclopean Perception* Chicago Press, Chicago.

Longuet-Higgins H. C. (1982a) The role of the vertical dimension in stereoscopic vision. *Perception* **11**, 337–386.

Longuet-Higgins H. C. (1982b) Appendix to paper by John Mayhew entitled: "The interpretation of stereo-disparity information: the computation of surface orientation and depth". *Perception* **11**, 405–407.

Marr D. (1982) *Vision*. Freeman, San Francisco, CA.

Mayhew J. (1982) The interpretation of stereo-disparity information: the computation of surface orientation and depth. *Perception* **11**, 387–403.

McKee S. P. (1983) The spatial requirements for fine stereoacuity. *Vision Res.* **23**, 191–198.

Mitchison G. J. and Westheimer G. (1984) The perception of depth in simple figures. *Vision Res.* **24**, 1063–1073.

Morrison J. D. and Whiteside T. C. D. (1984) Binocular cues in the perception of distance of a point source of light. *Perception* **13**, 555–566.

Ninio J. and Mizraji E. (1985) Errors in the stereoscopic separation of surfaces represented with regular textures. *Perception* **14**, 315–328.

Ono H. and Comerford J. (1977) Stereoscopic depth constancy. In *Stability and Constancy in Visual Perception: Mechanisms and Processes*. (Edited by W. Epstein). Wiley, New York.

Perkins D. N. (1972) Visual discrimination between rectangular and nonrectangular parallelopipeds. *Percept. Psychophys.* **12**, 396–400.

Prazdny K. (1983) Stereoscopic matching, eye position, and absolute depth. *Perception* **12**, 151–160.

Ritter M. (1977) Effect of disparity and viewing distance on perceived depth. *Percept. Psychophys.* **22**, 400–407.

Ritter M. (1979) Perception of depth: processing of simple positional disparity as a function of viewing distance. *Percept. Psychophys.* **25**, 209–214.

Rogers B. J. and Graham M. E. (1983) Anisotropies in the perception of three-dimensional surfaces. *Science, N.Y.* **221**, 1409–1411.

Schriever W. (1925) Experimentelle Studien uber das stereoskopische Sehen. *Z. Psychol.* **96**, 113–170.

Shepard R. N. (1981) Psychophysical Complementarity. In *Perceptual Organization* (Edited by M. Kubovy and J. Pomerantz). Lawrence Erlbaum, Hillsdale, N.J.

Smith A. H. (1965) Interaction of form and exposure time in the perception of slant. *Percept. Motor Skills* **20**, 481–490.

Smith O. W. and Smith P. C. (1957) Interaction of the effects of cues involved in judgements of curvature. *Am J. Psychol.* **70**, 361–375.

Smith P. C. and Smith O. W. (1961) Veridical perceptions of cylindricality: a problem of depth discrimination and object identification. *J. exp. Psychol.* **62**, 145–152.

Stevens K. A. (1981a) The visual interpretation of surface contours. *Artificial Intell.* **217**, 47–74.

Stevens K. A. (1981b) The information content of texture gradients. *Biol. Cybernet.* **42**, 95–105.

Stevens K. A. (1983a) Surface tilt (the direction of slant): a neglected psychophysical variable. *Percept. Psychophys.* **33**, 241–250.

Stevens K. A. (1983b) Slant-tilt: the visual encoding of surface orientation. *Biol. Cybernet.* **46**, 183–195.

Stevens K. A. (1984) On gradients and texture "gradients" Commentary on: Cutting & Millard 1984. Three gradients and the perception of flat and curved surfaces. *J. exp. Psychol. Gen.* **113**, 217–220.

Stevens K. A. and Brookes A. (1988) Theory of depth reconstruction in stereopsis. Submitted.

Tyler C. W. (1973) Stereoscopic vision: cortical limitations and a disparity scaling effect. *Science, N.Y.* **181**, 276–278.

Wallach H., Gillam B. and Cardillo L. (1979) Some consequences of stereoscopic depth constancy. *Percept. Psychophys.* **26**, 235–240.

Wallach H. and Bacon J. (1976) Two forms of retinal disparity. *Percept. Psychophys.* **19**, 375–382.

Wallach H. and Zuckerman C. (1963) The constancy of stereoscopic depth. *Am. J. Psychol.* **76**, 404–412.

Westheimer G. (1979) The spatial sense of the eye. *Invest. Ophthal. visual Sci.* **18**, 893–912.

Wheatstone C. (1852) On some remarkable, and hitherto unobserved phenomena of binocular vision. *Phil. Mag Ser.* **4**, 504–523.

Yellott J. I. and Kaiwi J. L. (1979) Depth inversion despite stereopsis: the appearance of random-dot stereograms on surfaces seen in reverse disparity. *Perception* **8**, 135–142.

Youngs W. M. (1976) The influence of perspective and disparity cues on the perception of slant. *Vision Res.* **16**, 79–82.

OFFICE OF NAVAL RESEARCH
PERCEPTUAL SCIENCE PROGRAMS -CODE 1142PS
TECHNICAL REPORTS DISTRIBUTION LIST

Dr. Earl Alluisi
Office of the Deputy Under Secretary of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D.C. 20301

Naval Aerospace Medical Research Laboratory
Sensory Division Code 23
Pensacola, FL 32508

Aircrew Systems Branch
Systems Engineering Test Directorate
U.S. Naval Test Center
Patuxent River, MD 20670

CDR R. Carter
Office of Chief on Naval Operations
OP-933D3
Washington, D.C. 20350

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 North Quincy Street
Arlington, VA 22217-5000

Commander
Naval Air Systems Command
Crew Station Design
NAVAIR 5313
Washington, D.C, 20361

Commanding Officer
Naval Health Research Center
San Diego, CA 92152

Dean of the Academic Departments
U.S. Naval Academy
Annapolis, MD 21402

Director
Technical Information Division
Code 2627
Naval Research Laboratory
Washington, D.C. 20375-5000

Dr. Robert A. Fleming
Human Factors Support Group
Naval Personnel Research and Development Center
1411 South Fern Street
Arlington, VA 22202-2896

Mr. Jeff Grossman
Human Factors Laboratory, Code 71
Navy Personnel R&D Center
San Diego, CA 92152-6800

Mr. Paul Heckman
Naval Ocean Systems Center
San Diego, CA 92152

CAPT William M. Houk
Commanding Officer
Naval Medical R&D Command
Bethesda, MD 20814-5055

Human Factors Branch
Code 3152
Naval Weapons Center
China Lake, CA 93555

Human Factors Department
Code N-71
Naval Training Systems Center
Orlando. FL 32813

Human Factors Engineering
Code 441
Naval Ocean Systems Center
San Diego, CA 92152

Code 222
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Michael Letsky
Office of the Chief of Naval Operations
(OP-01B7)
Washington, D.C. 20350

Lt Dennis McBride
Human Factors Branch
Pacific Missile Test Center
Point Mugu, CA 93042

CAPT W. Moroney
Naval Air Development Center
Code 602
Warminster, PA 18974

Perceptual Science Program
Office of Naval Research
Code 1142PS
800 North Quincy Street
Arlington, VA 22217-5000

LCDR T. Singer
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

Special Assistant for Marine Corps
Matters
Code 00MC
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Michael Drillings
Basic Research Office
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr.. Edgar M. Johnson
Technical Director
U.S. Army Research Institute
Alexandria, VA 22333-5600

Technical Director
U.S. Army Human Engineering Laboratory
Aberdeen Proving Ground, MD 21005

Mr. Charles Bates, Director
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Dr. Kenneth R. Boff
AF AMRL/HE
Wright-Patterson AFB, OH 45433

Dr. A. Fregly
Life Sciences Directorate, AFOSR
Bolling AFB, Bldg 410
Washington, D.C. 20332-6448

Defense Technical Information Center
Cameron Station, Bldg. 5
Alexandria, VA 22314

Dr. Alan Leshner
Division of Behavioral and Neural Sciences
Nation Science Foundation
1800 G. Street, N.W.
Washington, D.C. 20550

Dr. H. E. Bamford
Program Director
Division of Information, Robotics and Intelligent Systems
National Science Foundation
Washington, D.C. 20550

Dr. Donald D. Hoffman
University of California (Irvine)
School of Social Sciences
Irvine, CA 92717

Dr. Robert A. Hummel
New York University
Courant Inst. of Mathematical Sciences
251 Mercer Street
New York, New York 10012

Dr. David L. Kleinman
Electrical Engineering and Computer Science Department
University of Connecticut
Storrs, CT 06268

Dr. Alexander Levis
Massachusetts Institute of Technology
Lab Information & Decision Systems
Cambridge, MA 02139

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311

Dr. Tomaso Poggio
Massachusetts Institute of Technology
Center for Biological Information Processing
Cambridge, MA 02139

Dr. William B. Rouse
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, BA 30332

Dr. T. B. Sheridan
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

Mr. Demosthenis Teneketzis
Department of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI 48109-1109

Dr. Douglas Towne
University of Southern California
Behavioral Technology Laboratory
1945 South Elena Avenue, Fourth Floor
Redondo Beach, CA 90277

Dr. H. P. Vann Cott
NAS-National Research Council
(COHF)
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Dr. Christopher Wickens
Department of Psychology
University of Illinois
Urbana, IL 61801